PERCEIVING ACOUSTIC SOURCE ORIENTATION IN THREE-DIMENSIONAL SPACE

John G. Neuhoff

The College of Wooster
Department of Psychology
Wooster, OH 44691, U.S.A.
jneuhoff@wooster.edu
http://pages.wooster.edu/jneuhoff/index.htm

ABSTRACT

In virtual environments and auditory displays accurate representation of the simulated location and relative distance of a sound source can enhance the effectiveness of the display. However, in addition to having a physical location, many sound sources also project sound directionally. The ability of listeners to determine the facing direction (or facing angle) of unidirectional sound sources has been studied very little. In two experiments listeners estimated the facing angle of a unidirectional loudspeaker. In Experiment 1 listeners estimated the static directional orientation while blindfolded. In Experiment 2 listeners were given dynamic rotational cues prior to making judgments of orientation. The results show a significant advantage in perceiving facing angle when dynamic cues are present. Listeners were also significantly better at perceiving loudspeaker facing angle when closer to the source and when the loudspeaker was directly facing the listener.

1. INTRODUCTION

The use of sound in virtual environments has served as an impetus for research on spatial hearing. Much of this research has focused on horizontal or vertical localization and auditory distance perception. Although the spatial resolution of the auditory system does not match that of the visual system, it is nonetheless an important component in producing realistic virtual environments [1]. In addition to estimates of physical location, listeners are able to make reasonable estimates of many other physical characteristics of sound sources. For example, listeners can use acoustic cues to discriminate object length [2], shape [3, 4], and even use higher order temporal properties to perceive and categorize dynamic events such as breaking, bouncing, and vessel filling [5, 6]. Listeners can also determine whether there is room to pass between a sound source and a barrier [7], and whether there is an occluding object between a sound source and the listener [8, 9].

All of these abilities can be important in virtual environments and auditory displays in which listeners must perform navigation tasks and identify, localize, or manipulate virtual objects. However, another characteristic of many acoustic sources has been virtually ignored in both the psychoacoustic and auditory display literature. Many sound sources (e.g., talkers and loudspeakers) have a directional component such that they project sound in one primary hemifield. In this paper, this directional component is referred to as the "facing angle" of an acoustic source. The facing angle is formed by a line between a source and a listener, and a ray in the direction in which the source is radiating. Although intuitively, it seems that listeners might be sensitive to the facing angle of sources, there is virtually no empirical data on this ability or precision with which this task might be accomplished. Furthermore, there are several potential acoustic cues that listeners may use to perceive a source's facing angle. None of these cues have been systematically examined in the context of perceiving the facing angle of acoustic sources.

One cue to facing angle that is available to listeners results from the directivity of the acoustic source. The directivity characteristics of a loudspeaker can be obtained by measuring levels directly in front of the source and at equidistant angles around the source or by simply rotating the source while taking measurements from one position [10]. Directivity measurements for enclosed loudspeakers typically show peak levels directly in front of the source that drop off as the measurement point departs from 0^o. High frequency sounds are particularly directional and show greater "beaming" than low frequency sounds [10]. If listeners are able to discriminate the facing angle of a directional source, they may in part bas their judgments on interaural level differences (ILDs) that are created by the interaction of facing angle and the directivity of the source.

For example, it is often stated that ILDs do not exist when a sound source is in the median plane of the listener. However, directional sources in the median plane fail to produce ILDs only if the source is *directly* facing the listener. A directional source in the median plane *can* produce ILDs if it does not directly face the listener because of directivity characteristics. If the level measured directly in front of a source is higher than that measured at 10^o, then we might expect ILDs when the source is directly in the median plane but facing 10^o to one side of the listener. Thus, listeners may be able to use the ILD created by a source in the median plane as a cue to facing angle.

Another potential cue to facing angle is the ratio of direct to reflected sound that arrives at the listener. As the facing angle of a source departs from 0^o this ratio decreases. Listeners may be able to use this ratio and other characteristics of reflection to perceive facing angle in a manner similar to that used to judge auditory distance [11, 12].

Both the directivity/ILD and the ratio of direct-to-reflected sound cues predict that listeners will be able to more accurately perceive facing angles when closer to the acoustic source. Thus, the ability to determine the facing angle of a loudspeaker was examined from two different distances. Also of interest was the effect of dynamic rotation cues. Thus we examined accuracy of perceiving facing angle under conditions in which the loudspeaker was rotated while sounding (dynamic rotational cues), versus conditions in which the rotated loudspeaker sounded only after the rotation had stopped (static directional cues). To summarize, blindfolded listeners estimated the facing angle of a nearby loudspeaker after the speaker was rotated. In two separate experiments listeners performed this task with static directional cues and dynamic rotational cues from two different distances.

2. EXPERIMENT 1

2.1. Method

2.1.1. Participants

Thirty undergraduate students between the ages of 18 and 25 yrs. served as participants. All listeners reported normal hearing and received class credit for participation.

2.1.2. Apparatus and Stimuli

The experiment took place in a 2.74 m x 3.66 m room with 3 painted gypsum sheetrock walls and one painted concrete block wall, a 2.44 m high acoustical tile ceiling, and a carpeted floor. Stimuli were presented with a Koss portable CD player (Model HG 900). A male voice counting *"one, two, three, four, one, two, three four "* emanated from a Radio Shack Optimus XTS 40 loudspeaker at approximately 65 dB-A. Directivity measurements for the loudspeaker at three frequencies are shown in Figure 1. Stimulus duration was 4 s with one digit voiced every .5 s. Loudspeaker dimensions were 12.5 x 12.5 x 11.4 cm (HWD). The loudspeaker had a frequency response of 150-18,000 Hz and rested on a 91.4 x 61 cm table from which a 3 cm steel dowel 6.25 mm in diameter protruded. A 6.25 mm hole was drilled in the center of the bottom of the loudspeaker, and the speaker was placed on the table over the protruding steel dowel, thus allowing it to rotate freely in any direction (see Figure 2). The tabletop

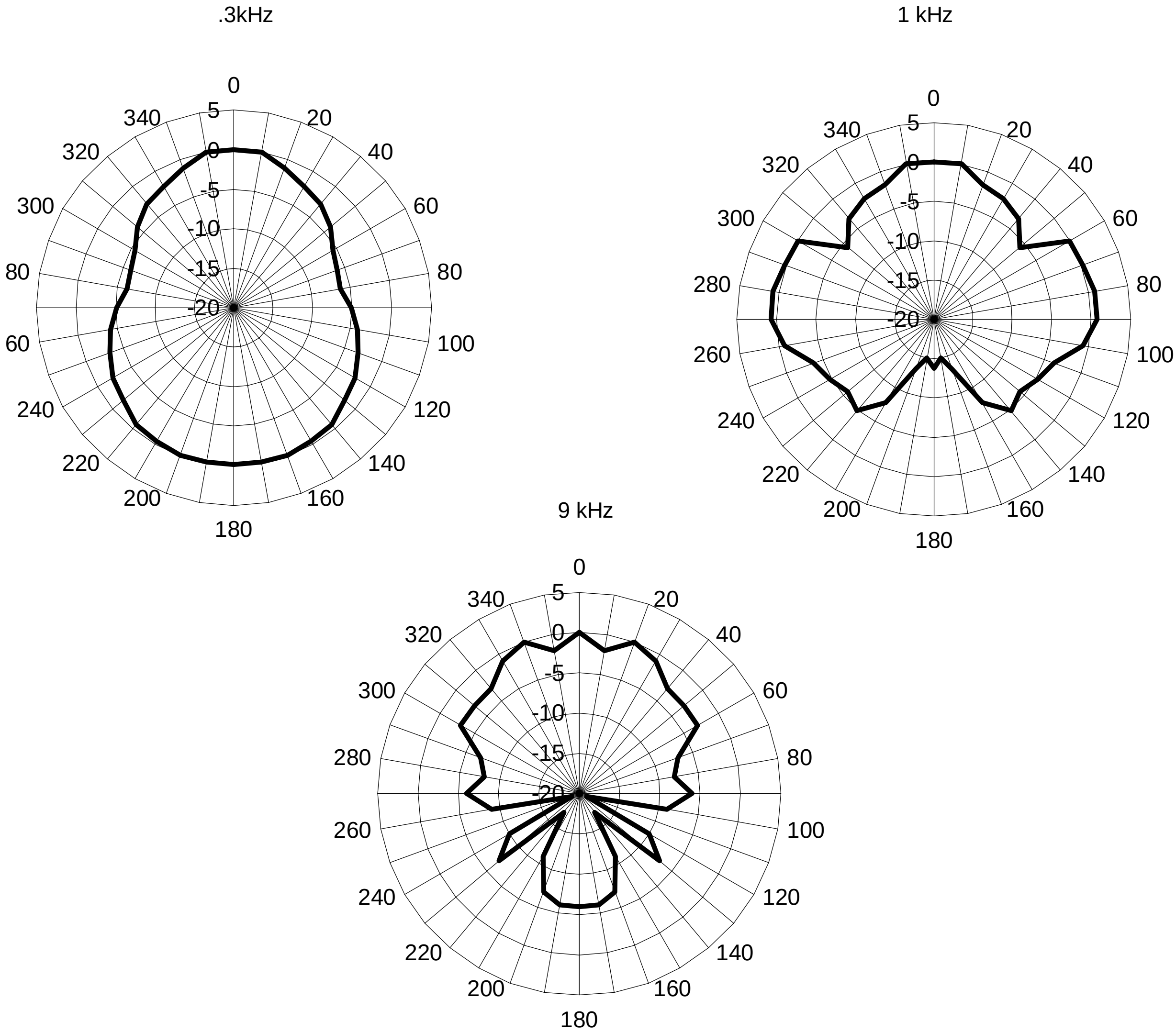

Figure 1. *Directivity patterns of the loudspeaker at .3, 1, and 9 kHz measured 1 m from the source in a reverberant room. Circumference is in degrees of facing angle. Radial axis is in dB.*

was 77.5 cm from the floor. Participants were seated at a 91.4 x 61 cm response table, the surface of which was 72.4 cm from the floor. A 3 cm steel dowel 6.25 mm in diameter also protruded from the center of the response table. A hole was drilled in the center of the bottom of a second Optimus XTS 40 loudspeaker, and the loudspeaker was placed over the protruding dowel in the response table so that it too was free to rotate in any direction. Both the stimulus and response speakers were fitted with a flat plastic pointer 6" in length that indicated the facing angle of each speaker by pointing to marks on the table surrounding each speaker. The response table was moved between blocks of trials so that, in two separate conditions, the distance between the two loudspeakers was .91 m and 1.82 m respectively.

2.1.3. Design and Procedure

Participants entered the experimental room and were seated at the response table. They were then blindfolded and told that they would hear a voice emanating from the stimulus loudspeaker and that the loudspeaker could be facing any direction. The listener's task was to indicate the facing angle of the stimulus loudspeaker by rotating the response loudspeaker to match the perceived facing direction of the stimulus loudspeaker. There were two trials from each of eight facing angles (0^o, 45^o, 90^o, 135^o, 180^o, 225^o, 270^o, 315^o). This provided for a total of sixteen randomly presented trials at each of the two listening distances. Half of the listeners provided responses from the .91 m listening distance first; the other half provided responses from the 1.82 m listening distance first. Prior to beginning experimental trials listeners were given two familiarization trials in which they were exposed to the stimulus sound and the acoustical properties of the room. On familiarization trials listeners heard the stimulus two times in succession while the speaker was rotated 360^o starting and ending at 0^o (facing the listener). On one familiarization trial the direction of rotation was clockwise; on the other trial the direction of rotation was counter-clockwise.

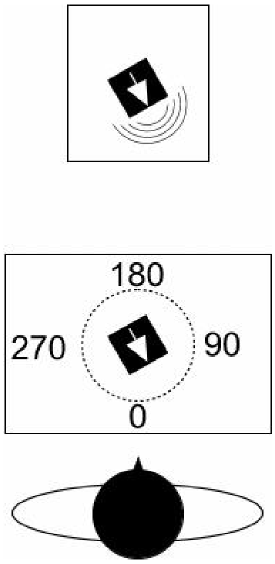

Figure 2. *Experimental setting. Blindfolded listeners adjusted the orientation of a response loudspeaker to match that of the stimulus loudspeaker.*

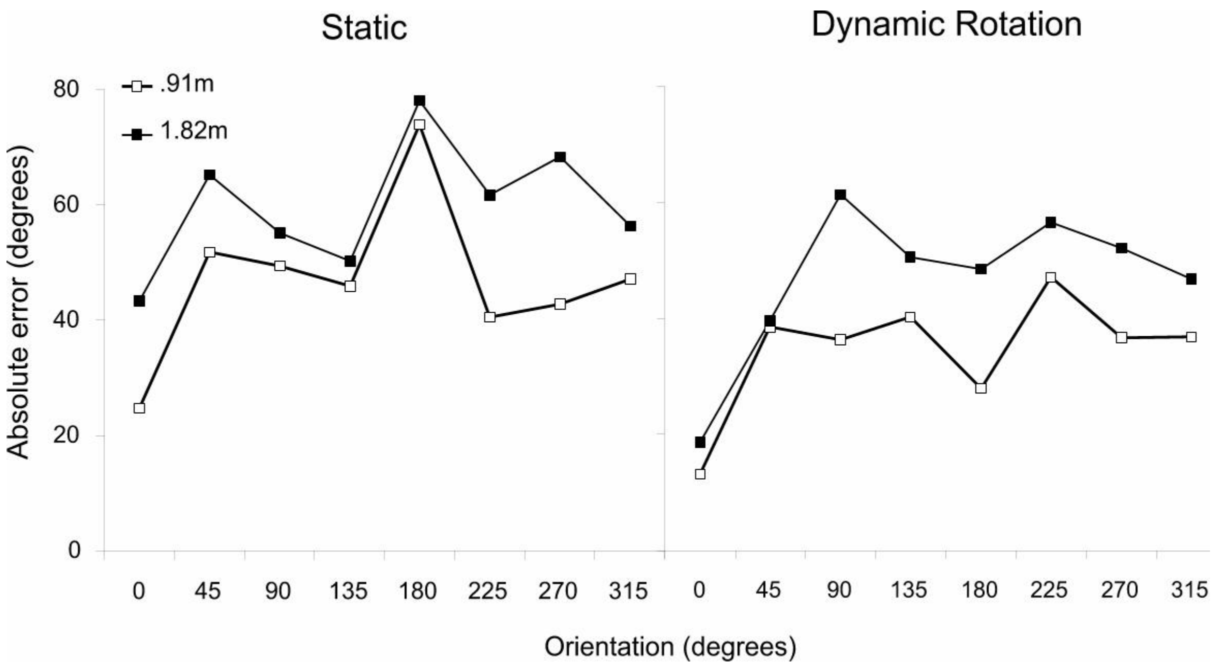

Figure 3. *Results of Experiments 1 (Static) and 2 (Dynamic Rotation). Performance is significantly better when listeners have dynamic rotation cues, are closer to the source, and when the loudspeaker is oriented at 0^o*

Listeners indicated their response by rotating the response speaker to the desired orientation and removing their hand from the speaker. The experimenter then recorded the facing angle of the response speaker.

2.2. Results

Each listener made two estimates at each of 8 facing angles. Estimates were averaged to obtain a single score at each orientation. A mean error score for each condition was calculated by taking absolute value of the difference between the perceived and the actual facing angle. This difference was constrained such that no error score could exceed 180^o. The mean error scores in each condition are show in Figure 3. A 2 (distance) x 8 (facing angle) analysis of variance (ANOVA) on the error scores showed a significant effect for facing angle, $\underline{F}$ (203, 7) = 3.09, p = .004. Listeners showed the best performance when the loudspeaker was oriented at 0^o (directly facing the listener). There was no significant effect for listening distance, $\underline{F}$ (29, 1) = 1.31, p = N.S.). However, there was a significant interaction between listening distance and facing angle , $\underline{F}$ (203, 7) = 7.11, p < .001.

Error scores that were between 165^o and 180^o were defined as "reversals". Averaged across all conditions only 4.6% of the trials were reversals. However, the majority of these occurred at 180^o with listeners mistaking the 180^o facing angle for 0^o. A chi-square test showed a significant difference in the number of reversals between facing angles, χ^2 (7) = 125.48, p < .001 (see Figure 4).

3. EXPERIMENT 2

3.1. Method

3.1.1. Participants

Twenty undergraduate students between the ages of 18 and 25 yrs. Served as participants. All listeners reported normal hearing and received class credit for participation.

3.1.2. Apparatus and Stimuli

The apparatus and stimuli used in Experiment 2 were identical to that used in Experiment 1.

3.1.3. Design and Procedure

Participants entered the experimental room and were seated at the response table. They were then blindfolded and told that they would hear a voice emanating from the stimulus loudspeaker. They were also told that the loudspeaker would be rotated when the voice began. The listener's task was to indicate the terminal trajectory of the stimulus loudspeaker by rotating the response loudspeaker to match the perceived facing direction of the stimulus loudspeaker. Prior to each trial, the facing direction of the response loudspeaker was aligned with that of the stimulus loudspeaker. Listeners were instructed to feel the response loudspeaker prior to each trial so that they knew the starting orientation of the stimulus loudspeaker. There were two trials from each of eight starting trajectories (0^o, 45^o, 90^o, 135^o, 180^o, 225^o, 270^o, 315^o). On each trial the loudspeaker was rotated 180^o, once in the clockwise direction and once in the counterclockwise direction. The eight starting trajectories and two rotation directions provided for a total of sixteen randomly presented trials at each of the two listening distances. Half of the listeners provided responses from the .91 m listening distance first; the other half provided responses from the 1.82 m listening distance first. The rotation speed on each trial was approximately 90^o/s. Prior to beginning experimental trials listeners were given two familiarization trials in which they were exposed to the stimulus sound, the speaker rotation, and the acoustical properties of the room. On familiarization trials listeners heard the stimulus two times in succession while the speaker was rotated 360^o starting and ending at 0^o (facing the listener). On one familiarization trial the direction of rotation

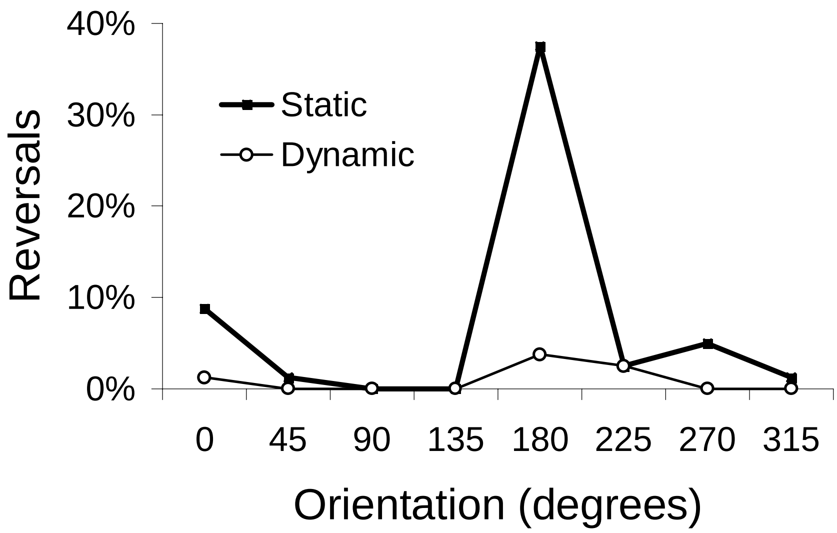

Figure 4. *Percentage of "reversals" (errors greater than 165^o) in facing angle estimates as a function of facing angle and rotation condition.*

was clockwise; on the other trial the direction of rotation was counter-clockwise. Listeners indicated their response by rotating the response speaker to the desired orientation and removing their hand from the speaker. The experimenter then recorded the facing angle of the response speaker.

3.2. Results

The difference between perceived facing direction and actual facing direction was calculated for each trial. These scores were converted to absolute vales, and mean error values were calculated in each condition. A 2 (listening distance) x 2 (rotation direction) x 8 (facing angle) repeated measures ANOVA was performed. Errors in perceived facing angle were significantly affected by actual facing angle, $\underline{F}$ (133, 7) = 5.3, $p < .001$. This effect appears to stem from greater accuracy when the speaker was oriented at 0^o (directly facing the listener, see Figure 2). When perceived facing angle was examined with 0^o removed, there was no significant difference between angles. Listeners were also significantly more accurate at estimating facing angle when they were closer to the source, $\underline{F}$ (19, 1) = 8.02, p = .011 (mean error .91m= 34^o, 1.82m = 47^o). Finally, there was a significant interaction between rotation direction (clockwise-counterclockwise) and facing angle, $\underline{F}$ (133, 7) = 3.19, p = .004.

To examine this interaction, the type of rotation was divided into two groups, those trials on which the speaker at some point in its path of rotation directly faced the observer and those trials on which it did not. For example, a clockwise trial that began at 90^o and ended at 270^o would rotate through 0^o, directly facing the listener at the midpoint of the rotation. However, the same angular rotation in the counterclockwise direction would rotate through 180^o and would not face the listener at any point in its rotation path. Excluded from this analysis were trials that ended at 0^o and 180^o because they did not pass through 0^o. Thus, a 2 (path) x 2 (distance) x 6 (angle) ANOVA was performed. The results showed that listeners were significantly more accurate in determining facing direction when the loudspeaker rotated toward them through 0^o than when it rotated away from them despite identical terminal orientation, F (19, 1) = 22.77, $p < .001$, mean toward = 39^o, away = 49^o). There was no significant difference for facing angle when both 0^o and 180^o were removed from the analysis, F (95, 5) = .75, p= N.S. However, listeners were still more accurate at the closer the listening distance than the farther, F (19, 1) = 7.67, p = .012 (mean error .91m= 37^o, 1.82m = 51^o).

The proportion of reversals (errors greater than 165 ^{o}) in Experiment 2 was less than 1%. A chi–square analysis failed to show a significant difference in the number of reversals across the eight different terminal orientations, χ^2 (7) = 12.67, p = N.S.). An analysis of reversals across Experiments 1 and 2 showed significantly more reversals in the static condition employed in Experiment 1 than in the dynamic condition employed in Experiment 2, χ^2 (1) = 76.92, $p < .001$ (see Figure 4).

4. DISCUSSION

Listeners made relatively accurate estimates of loudspeaker facing angle and showed a significant advantage when dynamic cues were available. In almost all conditions (excluding those where the source faced away from the listener) errors were almost always less than 60^o. Thus, listeners appeared to have a good sense of the "general direction" that the loudspeaker was facing. Performance was particularly good when the loudspeaker faced the listener directly and when the listener was closer to the source. The enhanced ability of listeners to localize egocentric orientation (0^o) may be responsible for the better accuracy at other facing angles when the source was rotated through 0^o. This may have provided listeners with an anchoring point from which they could better estimate other terminal facing angles.

Listeners were also better at determining facing angle when they were closer to the loudspeaker. At closer listening distances the ratio of direct to reflected sound is higher. Thus, this finding is consistent with the interpretation that listeners may in part rely on change in this ratio in making determinations of auditory source orientation for directional acoustic sources. If so, facing angle estimates based on dynamic rotation cues may rely on a tau-like function for change in the ratio of direct to reflected sound, similar to those suggested for intensity change in object approach [13] and frequency change in vessel filling [5].

Better performance at closer distances is also consistent with the hypothesis that listeners can use interaural level differences (ILD) in perceiving facing angle. For example, the directivity characteristics of our loudspeaker were such that levels were generally attenuated as the facing angle departed from 0^o. At 0^o then we would expect zero ILD. However as the loudspeaker was turned, the directivity pattern created ILDs that listeners may have used to perceive facing angle. At closer distances the amount of rotation required to create ILDs is smaller than that required at farther distances. Thus, our findings of greater precision at closer listening distances is also consistent with ILD as a cue to perceiving facing angle.

The perception of facing angle may be of particular importance with speech. Human listeners tend to visually orient toward the source of speech as well as project speech directionally toward the intended recipient of the message [14-16]. The fact that listeners are sensitive to facing angle suggests that incorporating facing angle in virtual environments might enhance intelligibility and more realistically approximate face-to-face communication, and ideal toward which many virtual communication systems strive [17].

5. REFERENCES

[1] R. H. Gilkey and T. R. Anderson, "Binaural and spatial hearing in real and virtual environments,". Mahwah, NJ: Lawrence Erlbaum, 1997.

[2] C. Carello, K. L. Anderson, and A. J. Kunkler-Peck, "Perception of object length by sound," *Psychological Science*, vol. 9, pp. 211-214, 1998.

[3] A. J. Kunkler-Peck and M. T. Turvey, "Hearing shape.," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 26, pp. 279-294, 2000.

[4] S. Lakatos, S. McAdams, and R. Causse, "The representation of auditory source characteristics: Simple geometric form," *Perception & Psychophysics*, vol. 59, pp. 1180-1190, 1997.

[5] P. A. Cabe and J. B. Pittenger, "Human sensitivity to acoustic information from vessel filling," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 26, pp. 313-324, 2000.

[6] W. H. Warren and R. R. Verbrugge, "Auditory perception of breaking and bouncing events: A case study in ecological acoustics," *Journal of Experimental Psychology: Human Perception and Performance*, vol. Vol 10, pp. 704-712, 1984.

[7] M. K. Russell and M. T. Turvey, "Auditory perception of unimpeded passage," *Ecological Psychology*, vol. Vol 11, pp. 175-188, 1999.

[8] M. K. Russell, "Acoustic perception of sound source occlusion," in *Studies in Perception and Action IV: Ninth International Conference on Perception and Action*, M. A. Schmuckler and J. M. Kennedy, Eds. Mahwah, NJ: Erlbaum, 1997, pp. 92-94.

[9] H. Ader, "Ein neues Hoerphaenomen.," *Monatsschrift fuer Ohrenheilkunde*, pp. 7, 1935.

[10] L. L. Beranek, *Acoustical Measurements*: American Institue of Physics, 1993.

[11] A. W. Bronkhorst and T. Houtgast, "Auditory distance perception in rooms," *Nature*, vol. 397, pp. 517-520, 1999.

[12] P. Zahorik and F. L. Wightman, "Loudness constancy with varying sound source distance," *Nature Neuroscience*, vol. 4, pp. 78-83, 2001.

[13] B. K. Shaw, R. S. McGowan, and M. T. Turvey, "An acoustic variable specifying time to contact," *Ecological Psychology*, vol. 3, pp. 253-261, 1991.

[14] P. Bertelson, J. Morais, P. Mousty, and C. Hublet, "Spatial constraints on attention to speech in the blind," *Brain & Language*, vol. 32, pp. 68-75, 1987.

[15] C. H. Brown, "The measurement of vocal amplitude and vocal radiation pattern in blue monkeys and grey-cheeked mangabeys," *Bioacoustics*, vol. 1, pp. 253-271, 1989.

[16] L. Ecklund Flores and G. Turkewitz, "Asymmetric headturning to speech and nonspeech in human newborns," *Developmental Psychobiology*, vol. Vol 29, pp. 205-217, 1996.

[17] M. T. Palmer, "Interpersonal communication and virtual reality: Mediating interpersonal relationships." in *Communication in the age of virtual reality., LEA's communication series.*, F. Biocca, (Ed) and M. R. Levy, (Ed), Eds.: Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US 1995, pp. 277-299.